

CLAIMS

I Claim:

1. A method of dynamically adjusting a time frame of a reservation in a compute environment, the method comprising:
receiving from a requestor a request for a reservation of resources in the cluster environment;
evaluating resources within the cluster environment to determine if a time frame for a reservation can be improved; and
if the time frame for the reservation can be improved, then migrating the reservation from the first group of resources at one time to a second group of resources at a second time.
2. The method of claim 1, wherein the cluster environment is one of an enterprise compute farm, a cluster and a grid.
3. The method of claim 1, wherein the reservation is used to guarantee a response time of a job according to cluster environment attributes.
4. The method of claim 1, wherein the time frame for a reservation further comprises one of the earliest possible time to start a reservation, an exact time to start a reservation, and a time duration of a reservation.
5. The method of claim 4, wherein the improved time frame for a reservation is one of an expanded duration time for the reservation or a contracted duration of time for the reservation.
6. The method of claim 1, wherein the improved time frame for a reservation comprises optimizing a response time for jobs submitted within the reservation.
7. The method of claim 1, further comprising dynamically modifying the second group of resources according to the requested reservation.
8. The method of claim 7, further comprising reserving the modified second group of resources.
9. The method of claim 7, wherein modifying the second group of resources further comprises modifying attributes of the second group of resources so as to meet requirements of the request.
10. The method of claim 1, wherein evaluating resources further comprises identifying accessible reserved resources and free resources.

11. The method of claim 10, wherein the accessible reserved resources are resources which are contained within a reservation to which the requestor has ownership of or priority access to.
12. The method of claim 1, wherein the first group of resources and the reservation for the second group of resources overlap in space and time.
13. The method of claim 1, wherein evaluating resources further comprises identifying resources that currently meet requestor criteria or which could meet requestor criteria through modifying the resources.
14. The method of claim 1, wherein the request comprises at least one of a requirement and a preference.
15. The method of claim 14, wherein the requirement and preference relate to at least one of a start time, a response time, optimization, quality of service, resource quantity, cost and time duration.
16. The method of claim 1, wherein the request comprises at least one request criteria of at least one of: computer operating system, software, network configuration information, file system configuration and memory requirements.
17. The method of claim 1, further comprising, after migrating the reservation, , providing an access list on the reservation so as to guarantee to the requestor a response time for submitted jobs.
18. The method of claim 17, wherein the access list comprises both credential-based data and performance or QOS-based data.
19. The method of claim 1, wherein the request for resources comprises a required criteria and a preferred criteria.
20. The method of claim 19, wherein the reservation of the first group of resources meets the required criteria.
21. The method of claim 20, wherein evaluating resources within the cluster environment to determine if the time frame can be improved further comprises evaluating resources to determine if at least one of the preferred criteria can be met by migrating the reservation in time.
22. The method of claim 21, wherein the determination of whether the time frame can be improved includes a comparison of a cost of migrating the reservation from the first group of

resources at the first time to the second group of resources at the second time with the improved time frame gained from meeting at least one of the preferred criteria.

23. The method of claim 22, wherein the cost comprises at least one of: provisioning nodes, dynamically allocating data and network access and allocating software licensing associated with customizing resources to meet the requirement of the requestor.

24. The method of claim 22, wherein the reservation of first group of resources is only migrated if the cost of migrating the reservation is less than the benefits of the improved time frame gained by meeting at least one of the preferred criteria.

25. The method of claim 1, wherein evaluating resources to improve the time frame is based on a per-reservation policy.

26. The method of claim 25, wherein the per-reservation policy is at least one of an administrator policy, a user-based policy, a policy of never taking an action, a policy of always taking an action and a cost-based policy.

27. The method of claim 1, wherein the requestor may identify the request as a self-optimizing request.

28. The method of claim 27, wherein a requestor is charged more for a self-optimizing request.

29. A system for dynamically adjusting a time frame of a reservation in a compute environment, the system comprising:

means for receiving from a requestor a request for a reservation of resources in the cluster environment;

means for evaluating resources within the cluster environment to determine if a time frame for a reservation can be improved; and

means for migrating the reservation from the first group of resources at one time to a second group of resources at a second time if the time frame for the reservation can be improved.

30. A system for dynamically adjusting a time frame of a reservation in a compute environment, the system comprising:

a module configured to receive from a requestor a request for a reservation of resources in the cluster environment;

a module configured to evaluate resources within the cluster environment to determine if a time frame for a reservation can be improved; and

a module configured to migrate the reservation from the first group of resources at one time to a second group of resources at a second time if the time frame for the reservation can be improved.

31. A computer-readable medium storing instructions for controlling a compute device to dynamically adjust a time frame of a reservation in a compute environment, the instructions comprising:

receiving from a requestor a request for a reservation of resources in the cluster environment;

evaluating resources within the cluster environment to determine if a time frame for a reservation can be improved; and

if the time frame for the reservation can be improved, then migrating the reservation from the first group of resources at one time to a second group of resources at a second time.